ImmPort FLOCK
Webex Demo Introduction

Richard H. Scheuermann, Ph.D.
U.T. Southwestern Medical Center

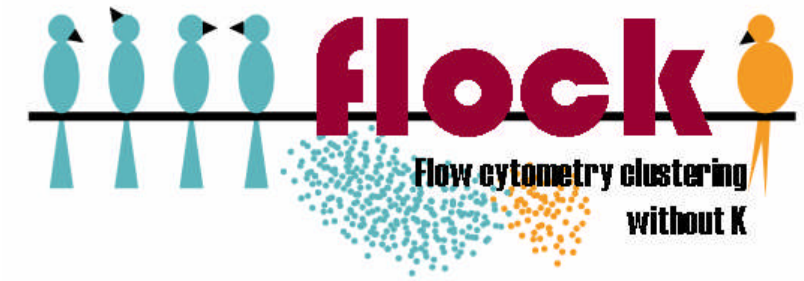
IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

Webex Objectives

- The mechanics of using ImmPort's FLOCK
- What FLOCK can and cannot do
- Future enhancements

IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

Flow Cytometry Analysis

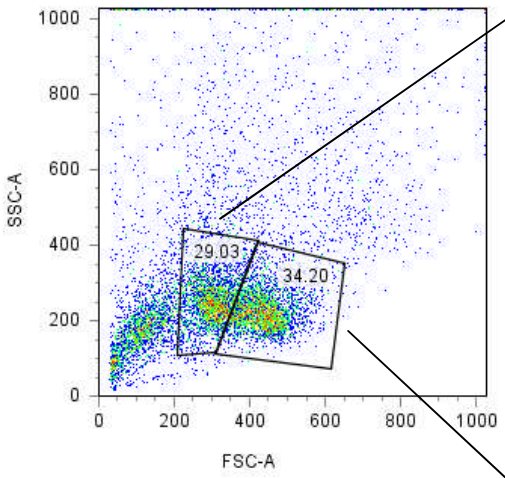


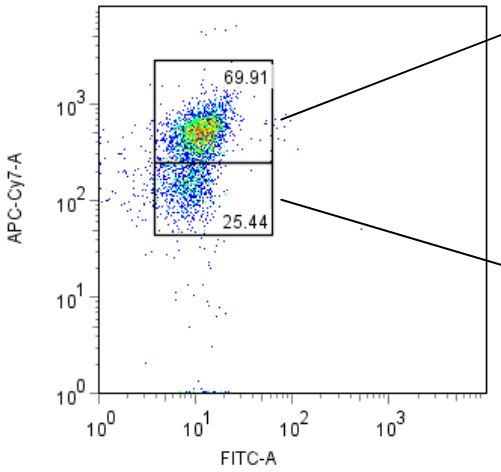
FLOCK is an algorithmic application for the identification of unique cell populations in multi-dimensional flow cytometry data

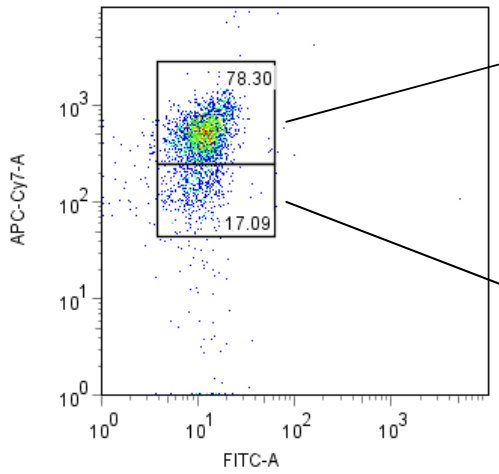
Traditional Flow Cytometry Analysis

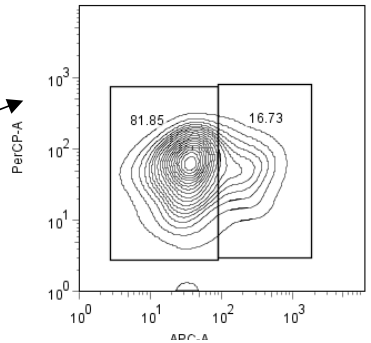
Goal - group together cells with similar characteristics

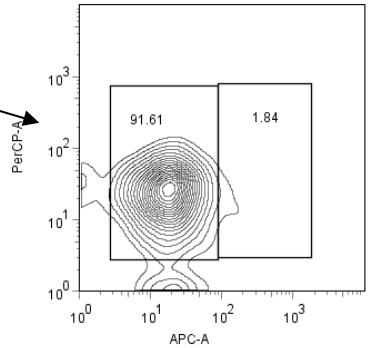
Traditional approach - manual gating 2D at a time

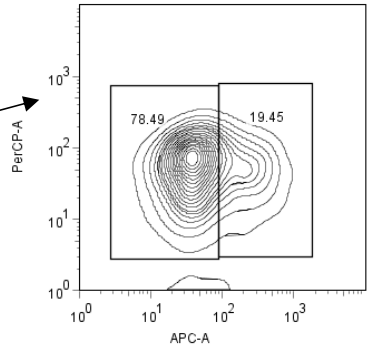


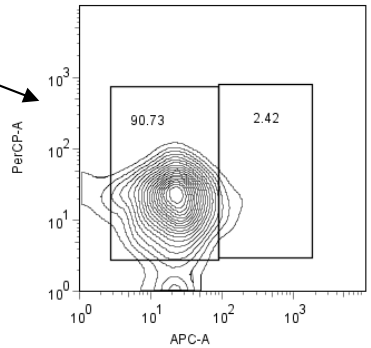












- Subjective
- Time-consuming
- Doesn't handle overlapping distributions well
- Sensitive to slight difference in fluorescence intensity distributions between samples
- Requires at least one 2D plot that clearly segregates populations in question

IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

New Approach

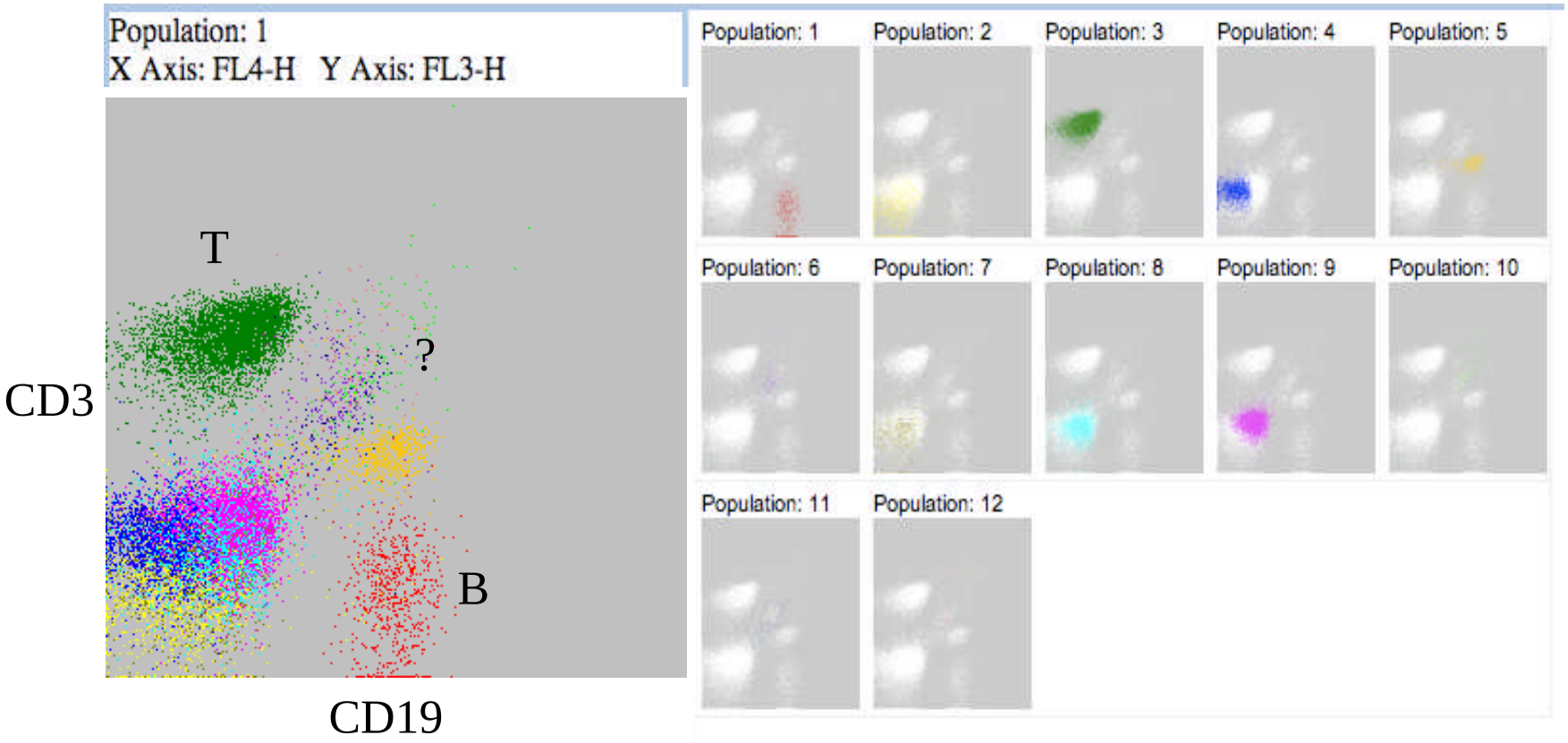
- Identifying cell populations automatically, objectively, and quickly in multi-dimensional flow cytometry data (eliminate manual gating)

- Quantitatively compare the identified populations across different samples and across different experiments

FLOCK v2.0 STEPS

1. ***File Conversion*** - Convert binary .fcs file into a data matrix
2. ***Data Cleansing*** - Remove boundary events (noise) in FSC and SSC dimensions
3. ***Data Shrinking*** - Collapse data toward distribution modes
4. ***Normalization*** - Z-score normalization for values in each dimension $((x_i - \mu)/SD)$
5. ***Dimension Selection*** - Select most informative dimensions based on measures of dispersion and distortion
6. ***FLOCK LoD*** – grid-based density clustering
   i. Partition each dimension to generate a hyper-grid
   ii. Identify dense hyper-regions in hyper-grid
   iii. Merge neighboring dense hyper-regions to define hyper-region groups (*n*)
   iv. Determine centroids for each hyper-region group
   v. Use *n* centroids to seed single round of distance-based clustering
7. ***FLOCK HiD*** - Refine population definition based on histogram partitioning
8. ***Group Merging*** - Merge close hyper-region groups based on [distance metric]
9. ***Centroid Calculation*** - Compute centroid for each hyper-region group
10. ***Clustering*** - Cluster events to nearest centroid
11. ***Population statistics*** - Summarize population proportions, intensity levels, etc.
12. ***Visualization***

FLOCK Result Example



FLOCK in ImmPort

- FLOCK has been implemented as a web-based application in ImmPort
- ImmPort is:
  - The Immunology Database and Analysis Portal
  - Funded by DAIT/NIAID/NIH
  - Web-based resource, accessible using any internet browser
  - Open to the general scientific research community
  - Free
  - Archive of research data from DAIT-funded investigators
  - Suite of analytical and data mining tools
  - Also contains a compendium of reference data from public resources for data mining
  - Analysis in private project workspaces

ImmPort
Bioinformatics For The Future of Immunology

Demo

- ImmPort registration/log in
- Upload formatted FCM result files
- Data set generation – sample grouping for batch analysis
- Single sample FLOCK
- Review of single sample FLOCK analysis
- Manual parameter setting on representative sample
- Centroid adjustment
- Cross sample comparison
- View cross sample analysis results

FLOCK in ImmPort

- What FLOCK does
  - Automated population identification in multi-dimensional FCM data
  - Mapping populations across samples
- What FLOCK doesn’t do
  - Traditional manual gating
- Future enhancements
  - Bulk upload
  - Improved auto parameter estimation
  - Population overlay display
  - Track adjustments
  - Output data in a GatingML format
  - Link FLOCK results to ontology of cell types